

Optimal Foraging of Renewable Resources

John J. Enright Emilio Frazzoli

Abstract—Consider a team of agents in the plane searching for and visiting target points that appear in a bounded environment according to a stochastic renewal process with a known absolutely continuous spatial distribution. Agents must detect targets with limited-range onboard sensors. It is desired to minimize the expected waiting time between the appearance of a target point, and the instant it is visited. When the sensing radius is small, the system time is dominated by time spent searching, and it is shown that the optimal policy requires the agents to search a region at a relative frequency proportional to the square root of its renewal rate. On the other hand, when targets appear frequently, the system time is dominated by time spent servicing known targets, and it is shown that the optimal policy requires the agents to service a region at a relative frequency proportional to the cube root of its renewal rate. Furthermore, the presented algorithms in this case recover the optimal performance achieved by agents with full information of the environment. Simulation results verify the theoretical performance of the algorithms.

I. INTRODUCTION

A very active research area today addresses the coordination of several mobile agents: groups of autonomous robots and large-scale mobile networks are being considered for a broad class of applications, ranging from environmental monitoring, to search and rescue operations, and national security. Wide-area surveillance is one of the prototypical missions for Uninhabited Aerial Vehicles (UAVs): low-altitude UAVs on such a mission must provide coverage of a region and investigate events of interest as they manifest themselves. In particular, we are interested in cases in which close-range information is required on targets detected by onboard sensors, and the UAVs must proceed to the locations to gather further on-site information.

We address a routing problem for a team of agents in the plane: target points appear over time in a bounded environment according to a stochastic renewal process with a known absolutely continuous spatial distribution. It is desired to stabilize the outstanding target queue and minimize the expected elapsed time between the appearance of a target point, and the instant it is visited (the system time). This is a formulation of the Dynamic Traveling Repairman Problem (DTRP), introduced in [1] and thoroughly developed in [2], [3]. Numerous algorithms are presented and analyzed in this series of seminal works. Furthermore, the property of *spatial bias* is studied. In particular, they analyze the problem and develop policies under the constraint that a target's expected waiting time must be independent of its location. In addition to combinatorial and convex optimization, many of the solutions rely heavily on results from the relatively mature fields of facility location, probability and queueing theory.

In an effort to address issues relevant in applications such as autonomous mobile robotics, this paper focuses on a variation of the DTRP. We place limitations on the information available to the vehicles and analyze the effect on the system's achievable performance. In particular, we consider the case in which vehicles are not aware of the location of targets as they appear, but rather must detect them using on-board sensors with a limited range.

The recent literature concerning problems of this class is vast. Some recent work on the routing of nonholonomic vehicles includes [4], [5]. Many of the new results for the small sensing radius case are applicable to coverage problems [6], [7], [8], in which the agents spread out with some sense of balance, or comb the environment efficiently. This case also has connections to the Persistent Area Denial (PAD) and area coverage problems [9], [10], [11], [12]. Other works are concerned with the generation of efficient cooperative strategies for several mobile agents to move through a certain number of given target points, possibly avoiding obstacles or threats [13], [14], [15], [16]. Trajectory efficiency in these cases is understood in terms of cost for the agents: in other words, efficient trajectories minimize the total path length, the time needed to complete the task, or the fuel/energy expenditure. A related problem has been investigated as the Weapon-Target Assignment (WTA) problem, in which mobile agents are allowed to team up in order to enhance the probability of a favorable outcome in a target engagement [17], [18]. In other works addressing UAV task-assignment, target locations are known and an assignment strategy is sought that maximizes the global success rate [19], [20]. Moreover, this work holds a place in the recent wave of investigation into the cooperative control of multi-agent systems [21], [22]. Other works addressing cooperative task completion by UAVs include [23], [24], [25].

Song and coworkers considered the problem of searching for a static object emitting intermittent stochastic signals under a limited sensing range, and analyze the performance of standard algorithms such as systematic sweep and random walks [26]. Due to the intermittent signals from the object, robots must perform a persistent search, thus making the work similar to ours. However, the authors assumed no prior information about the location of the target object is available; hence, their setting is equivalent to the assumption of a uniform spatial distribution. In our work, we explicitly consider non-uniform spatial distributions, which lead to different kinds of optimal policies. Mathew and Mezic presented an algorithm named Spectral Multiscale Coverage (SMC) to devise trajectories such that the spatial distribution of a patrol vehicle's position asymptotically matches a given function [27]. Similarly, Cannata and Sgorbissa describe an algorithm that solves what they call the Multirobot Controlled Frequency Coverage (MRCFC)

John J. Enright is with Kiva Systems, jenright@kivasystems.com.

Emilio Frazzoli is with the Laboratory for Information and Decision Systems at the Massachusetts Institute of Technology, frazzoli@mit.edu.

problem, in which a team of robots are required to repeatedly visit a set of predefined locations according to a specified frequency distribution [28]. We show that when attempting to minimize discovery time, the desired spatial distribution of the agent's position is dependent on, but not equivalent to the underlying spatial distribution of incidents it must find.

The main contributions of this paper are the following. 1) For the case of small sensing radius, we establish a new lower bound and 2) a new policy whose performance is tight with the lower bound. 3) In combination, these two contributions show that the optimal policy requires that agents search subregions of the environment at relative frequencies proportional to the square root of their renewal rates, i.e., searching is biased towards high density regions, but not as biased as the density function itself. 4) For the case of high target renewal rate, we present a new policy that works for agents with limited sensing capabilities. 5) Comparing the performance of our policy with a previously known lower bound, we show that—like the full-information case—the optimal policy requires that agents search subregions of the environment at relative frequencies proportional to the cube root of their renewal rates. 6) These results imply that the limited sensing capabilities do not adversely affect the optimal performance of the agents in this case, i.e., we recover the optimal performance of the full-information heavy load case. 7) Moreover, we provide scalable, decentralized strategies by which a multi-vehicle team can operate with the above mentioned algorithms, and retain optimal performance. Earlier versions of this work only consider a single agent and uniform spatial distribution [29] or a single agent and a known absolutely continuous spatial distribution, but without analysis of the case of high target renewal rate [30].

This paper is organized as follows. In Sec. II we formulate the DTRP with limited sensing and review known results. In Sections III we offer a lower bound for this new problem. In IV we present algorithms for the single agent, and compare their performance with lower bounds. In Sec. V we adapt our algorithms to the multiple vehicle setting. Sec. VI concludes the paper and notes possibilities for future research.

II. PROBLEM FORMULATIONS AND PREVIOUS RESULTS

In this section, we formally introduce the dynamic vehicle routing problem we wish to study, without the additional limitations on sensing or motion constraints. We also review results of well studied static vehicle routing problems, in which the vehicles have full information, and travel cost is simply Euclidean distance. The known performance limits for these problems serve as reference points for results found on the problem variations studied herein. They give insight as to how the new constraints affect the efficiency of the system.

Given a set $\mathcal{D}_n \subset \mathbb{R}^2$ of n points, the two-dimensional Euclidean Traveling Salesman Problem (ETSP) is the problem of finding the shortest tour (closed path) through all points in \mathcal{D}_n ; let $\text{ETSP}(\mathcal{D}_n)$ be the length of such a tour. Furthermore, we will make use of the following remarkable result.

Theorem 1 ([31], [32]): If the locations of the points in \mathcal{D}_n are independently and identically distributed (i.i.d.) with

compact support $\mathcal{Q} \subset \mathbb{R}^2$, then with probability one

$$\lim_{n \rightarrow \infty} \frac{\text{ETSP}(\mathcal{D}_n)}{\sqrt{n}} = \beta \int_{\mathcal{Q}} \sqrt{\varphi(q)} dq, \quad (1)$$

where $\beta > 0$ is a constant not depending on the distribution of the points and where φ is the density of the absolutely continuous part of the distribution of the points.

The current best estimate of the constant is $\beta \approx 0.7120$ [33], [34]. According to [35], if \mathcal{Q} is a “fairly compact and fairly convex” set in the plane, then (1) provides an adequate estimate of the optimal ETSP tour length for values of n as low as 15. Interestingly, the asymptotic cost of the ETSP for uniform point distributions is an upper bound on the asymptotic cost for general point distributions, as can be proved by applying Jensen's inequality to (1). In other words, if A is the area of set \mathcal{Q} ,

$$\lim_{n \rightarrow \infty} \frac{\text{ETSP}(\mathcal{D}_n)}{\sqrt{n}} = \beta \int_{\mathcal{Q}} \sqrt{\varphi(q)} dq \leq \beta \sqrt{A}$$

with probability one.

We will present algorithms that require online solutions of large ETSPs. In practice, these solutions are computed using heuristics such as Lin-Kernighan's [36] or approximation algorithms such as Christofides' [37]. If the algorithm used in practice guarantees a performance within a constant factor of the optimal, the effect on the performances of our algorithms can be modeled as a scaling of the constant β .

The following is a formulation of the *Dynamic Traveling Repairman Problem* (DTRP) [2], [38], [3]. Let $\mathcal{Q} \subset \mathbb{R}^2$ be a convex, compact domain on the plane, with non-empty interior; we will refer to \mathcal{Q} as the *environment*. Let A be the area of \mathcal{Q} . Target points are i.i.d. and generated according to a spatio-temporal Poisson point process, with temporal intensity $\lambda > 0$, and an absolutely continuous spatial distribution described by the density function $\varphi : \mathcal{Q} \rightarrow \mathbb{R}_+$. The spatial density function φ is K-Lipschitz, $|\varphi(q_1) - \varphi(q_2)| \leq K\|q_1 - q_2\|$ for all q_1 and q_2 in \mathcal{Q} , and bounded above and below, $0 < \underline{\varphi} \leq \varphi(q) \leq \bar{\varphi} < \infty$ for all q in \mathcal{Q} , and is normalized in such a way that $\int_{\mathcal{Q}} \varphi(q) dq = 1$. For any $t > 0$, $\mathcal{P}(t)$ is a random collection of points in \mathcal{Q} , representing the targets generated in the time interval $[0, t)$. The following are consequences of the properties defining Poisson processes.

- The total numbers of targets generated in two disjoint time-space regions are *independent* random variables.
- The total number of targets generated in an interval $[t, t + \Delta t)$ in a measurable set $\mathcal{S} \subseteq \mathcal{Q}$ satisfies

$$\begin{aligned} \Pr[\text{card}((\mathcal{P}(t + \Delta t) - \mathcal{P}(t)) \cap \mathcal{S}) = k] \\ = \frac{\exp(-\lambda \Delta t \cdot \varphi(\mathcal{S})) (\lambda \Delta t \cdot \varphi(\mathcal{S}))^k}{k!}, \end{aligned}$$

for all $k \in \mathbb{N}$, and hence

$$\mathbb{E}[\text{card}((\mathcal{P}(t + \Delta t) - \mathcal{P}(t)) \cap \mathcal{S})] = \lambda \Delta t \cdot \varphi(\mathcal{S}),$$

where $\varphi(\mathcal{S})$ is a shorthand for $\int_{\mathcal{S}} \varphi(q) dq$.

A service request is fulfilled when one of m mobile agents, modeled as point masses, moves to the target point associated with it; m is a possibly large, but finite number. In later sections, we will introduce limitations on the agent's awareness of

targets, and nonholonomic constraints on the agent's motion. Let us assume the agents have holonomic (single integrator) dynamics, and that all agent's are aware of a target's location upon its arrival epoch. Let $p(t) = (p_1(t), p_2(t), \dots, p_m(t)) \in \mathcal{Q}^m$ be a vector describing the positions of the agents at time t . The agents are free to move, with bounded speed, within the workspace \mathcal{Q} ; let v be the maximum speed of the agents. In other words, the dynamics of the agents are described by differential equations of the form

$$\frac{d p_i(t)}{dt} = u_i(t), \quad \text{with } \|u_i(t)\| \leq v, \quad \forall t \geq 0, i \in I_m,$$

where we denote $I_m = \{1, \dots, m\}$. The agents are identical, and have unlimited fuel and target-servicing capacity.

Let $\mathcal{D} : t \rightarrow 2^{\mathcal{Q}}$ indicate (a realization of) the stochastic process obtained combining the target generation process \mathcal{P} and the removal process caused by the agents visiting outstanding requests. The random set $\mathcal{D}(t) \subset \mathcal{Q}$ represents the *demand*, i.e., the service requests outstanding at time t ; let $n(t) = \text{card}(\mathcal{D}(t))$.

A motion coordination policy is a function that determines the actions of each vehicle over time, based on the locally-available information. For the time being, we will denote these functions as $\pi = (\pi_1, \pi_2, \dots, \pi_m)$, but do not explicitly state their domain; the output of these functions is a velocity command for each agent. Our objective is the design of motion coordination strategies that allow the mobile agents to fulfill service requests efficiently (we will make this more precise in the following).

A policy $\pi = (\pi_1, \pi_2, \dots, \pi_m)$ is said to be *stabilizing* if, under its effect, the expected number of outstanding targets does not diverge over time, i.e., if

$$\bar{n}_\pi = \lim_{t \rightarrow \infty} \mathbb{E}[n(t) : \text{agents execute policy } \pi] < \infty.$$

Intuitively, a policy is stabilizing if the mobile agents are able to visit targets at a rate that is—on average—at least as fast as the rate at which new service requests are generated.

Let T_j be the time elapsed between the generation of the j -th target, and the time it is fulfilled. If the system is stable, then the following balance equation (also known as Little's formula [39]) holds:

$$\bar{n}_\pi = \lambda \bar{T}_\pi, \quad (2)$$

where $\bar{T}_\pi := \lim_{j \rightarrow \infty} \mathbb{E}[T_j]$ is the system time under policy π , i.e., the expected time a target must wait before being fulfilled, given that the mobile agents follow the strategy defined by π . Note that the system time \bar{T}_π can be thought of as a measure of the quality of service collectively provided by the agents.

At this point we can finally state our problem: we wish to devise a policy that is (i) stabilizing, and (ii) yields a quality of service (system time) achieving, or approximating, the theoretical optimal performance given by

$$\bar{T}_{\text{opt}} = \inf_{\pi \text{ stabilizing}} \bar{T}_\pi.$$

In the following, we are interested in designing control policies that provide constant-factor approximations of the optimal achievable performance; a policy π is said to pro-

vide a constant-factor approximation of κ if $\bar{T}_\pi \leq \kappa \bar{T}_{\text{opt}}$. Furthermore, a policy is called *spatially unbiased* if, under its action, a target's expected waiting time is independent of its location and *spatially biased* otherwise. We shall investigate how this spatial constraint effects the achievable system time, i.e., we shall find lower bounds and develop algorithms within the class of spatially unbiased policies, and without. Moreover, we are interested in decentralized, scalable, adaptive control policies, that rely only on local exchange of information between neighboring vehicles, and do not require explicit knowledge of the global structure of the network.

The DTRP with general demand distribution is studied in [3], where the form of the optimal system time in heavy load is first derived. However, there remained a constant-factor gap between the lower and upper bounds. The coefficient of the lower bound was tightened from $(2/(3\sqrt{\pi}))^2$ to $\beta^2/2$ in [40]:

$$\lim_{\lambda \rightarrow \infty} \frac{\bar{T}_{\text{opt}}}{\lambda} = \frac{\beta^2}{2m^2v^2} \left(\int_{\mathcal{Q}} \sqrt{\varphi(q)} dq \right)^2, \quad (3)$$

within the class of spatially unbiased policies, and

$$\lim_{\lambda \rightarrow \infty} \frac{\bar{T}_{\text{opt}}}{\lambda} = \frac{\beta^2}{2m^2v^2} \left(\int_{\mathcal{Q}} \varphi(q)^{2/3} dq \right)^3, \quad (4)$$

within the class of spatially biased policies. As mentioned in [3],

$$A \geq \left(\int_{\mathcal{Q}} \sqrt{\varphi(q)} dq \right)^2 \geq \left(\int_{\mathcal{Q}} \varphi(q)^{2/3} dq \right)^3$$

with equality holding throughout if and only if $\varphi(q) = 1/A$ for all $q \in \mathcal{Q}$. In other words, uniform density is the worst possible, and any non-uniformity will strictly lower the optimal system time. This is analogous with the length of the stochastic ETSP, i.e., Eq. (1). Furthermore, the optimal system time for spatially biased policies is lower than that of spatially unbiased policies. This follows intuition as spatially unbiased waiting time is a constraint limiting the realm of available policies.

In addition to the above formulation of the DTRP, we add a constraint on the information available to the agents. Agents are not aware of a target's existence or location upon its arrival epoch. Rather, they must detect targets with limited-range onboard sensors, i.e., they must come within the local vicinity of the target. Let us call this variation the Limited Sensing DTRP. Formally, this means the set $\mathcal{D}(t)$ is in general not entirely known to all agents, due to the fact that the sensing range is limited. For the sake of simplicity, we will model the sensing region of an agent as a disk of radius r centered at the agent's position; indicate with

$$d(p_i) = \{q \in \mathbb{R}^2 : \|p_i - q\| \leq r\}$$

the sensing region for the i -th agent. Other shapes of the sensor footprint can be considered with minor modifications to our analysis, and affect the results at most by a constant. We will assume that the sensor footprint is small enough that it is contained within \mathcal{Q} for at least one position of the agent; moreover, we will be interested in analyzing the effect of a limited sensor range as $r \rightarrow 0^+$.

III. LOWER BOUNDS ON THE OPTIMAL SYSTEM TIME

Every target must be detected by an agent's sensor before it is serviced. Let T_j^{detect} be the time elapsed between the generation of the j -th target, and the time it is first detected by any agent's sensor. Consider an alternative scenario we call the *Detection DTRP*, in which targets are fulfilled at the instant they are first detected, and $\bar{T}_\pi^{\text{detect}} := \lim_{j \rightarrow \infty} \mathbb{E}[T_j^{\text{detect}}]$ is the *detection time* under policy π . For all targets, $T_j^{\text{detect}} \leq T_j$, and thus

$$\inf_{\pi \text{ stabilizing}} \bar{T}_\pi \geq \inf_{\pi \text{ stabilizing}} \bar{T}_\pi^{\text{detect}}.$$

In other words, a lower bound on the achievable detection time is also a lower bound on the achievable system time, \bar{T}_{opt} . We leverage this fact in the following theorem.

Theorem 2: The optimal system time for the DTRP with limited sensing satisfies

$$\lim_{r \rightarrow 0^+} \bar{T}_{\text{opt}} r \geq \frac{A}{4mv}, \quad (5)$$

within the class of spatially unbiased policies and

$$\lim_{r \rightarrow 0^+} \bar{T}_{\text{opt}} r \geq \frac{1}{4mv} \left(\int_{\mathcal{Q}} \sqrt{\varphi(q)} dq \right)^2, \quad (6)$$

within the class of spatially biased policies.

Proof: The probability that a target's location is within a sensor footprint at the time of arrival is bounded by

$$\Pr[q \in \cup_{i=1}^m d(p_i)] \leq \bar{\varphi} m \pi r^2.$$

In this case, $T_j^{\text{detect}} = 0$. However, for any fixed number of vehicles m and distribution φ ,

$$\lim_{r \rightarrow 0^+} \Pr[q \in \cup_{i=1}^m d(p_i)] \leq \lim_{r \rightarrow 0^+} \bar{\varphi} m \pi r^2 = 0,$$

and therefore,

$$\lim_{r \rightarrow 0^+} \Pr[q \notin \cup_{i=1}^m d(p_i)] = 1.$$

In this limit, from the perspective of a point $q \in \mathcal{Q}$, the actions of a given stabilizing policy π are described by the following (possibly random) sequence of variables: the lengths of the time intervals during which the point is not contained in any sensor footprint, $Y_k(q)$. In the following, we use the notation

$$\mathbb{E}[Y_\pi(q)] = \lim_{k \rightarrow \infty} \mathbb{E}[Y_k(q) : \text{agents execute policy } \pi].$$

Due to random incidence [35], [41], a target's detection time, conditioned upon its location, can be written as

$$\begin{aligned} \lim_{r \rightarrow 0^+} \mathbb{E}[T^{\text{detect}}|q] &= \lim_{r \rightarrow 0^+} \Pr[q \notin \cup_{i=1}^m d(p_i)] \cdot \frac{\mathbb{E}[Y_\pi^2(q)]}{2\mathbb{E}[Y_\pi(q)]} \\ &= \frac{\mathbb{E}[Y_\pi(q)]^2 + \text{Var}[Y_\pi(q)]}{2\mathbb{E}[Y_\pi(q)]} \\ &\geq \frac{1}{2} \mathbb{E}[Y_\pi(q)] \end{aligned}$$

where $\mathbb{E}[Y_\pi^2(q)]$ and $\text{Var}[Y_\pi(q)]$ are, respectively, the second moment and variance of the sequence $Y_k(q)$ under the actions of policy π . In other words, for fixed $\mathbb{E}[Y_\pi(q)]$, the system time is minimized if $\text{Var}[Y_\pi(q)] = 0$. This occurs under the

actions of a deterministic policy with exactly regular time intervals between searching location q .

Define $f(q)$ as the—time averaged—frequency at which point q is searched under the actions of a policy. Note that $f(q) = 1/\mathbb{E}[Y_\pi(q)]$, and so

$$\begin{aligned} \bar{T}^{\text{detect}} &= \int_{\mathcal{Q}} \varphi(q) \mathbb{E}[T_j^{\text{detect}}|q] dq \\ &\geq \frac{1}{2} \int_{\mathcal{Q}} \frac{\varphi(q)}{f(q)} dq. \end{aligned}$$

The m -vehicle system is capable of searching at a maximum rate of $2mvr$ (area per unit time), and so the average searching frequency is bounded by $\int_{\mathcal{Q}} f(q)/A dq \leq 2mvr/A$. Recalling $\bar{T}_{\text{opt}} \geq \bar{T}^{\text{detect}}$ we have the following minimization problem:

$$\begin{aligned} 2\bar{T}_{\text{opt}} &\geq \min_f \int_{\mathcal{Q}} \frac{\varphi(q)}{f(q)} dq \quad \text{subject to} \\ \int_{\mathcal{Q}} f(q) dq &\leq 2mvr \quad \text{and} \quad f(q) > 0. \end{aligned}$$

Since the objective function is convex in $f(q)$ and the constraints are linear, the above is an infinite-dimensional convex program. Relaxing the constraint with a multiplier, we arrive at the Lagrange dual:

$$\begin{aligned} 2\bar{T}_{\text{opt}} &\geq \min_{f(q)>0} \left[\int_{\mathcal{Q}} \frac{\varphi(q)}{f(q)} dq + \Gamma \left(\int_{\mathcal{Q}} f(q) dq - 2mvr \right) \right] \\ &= \int_{\mathcal{Q}} \min_{f(q)>0} \left[\frac{\varphi(q)}{f(q)} + \Gamma f(q) \right] dq - 2mvr\Gamma. \end{aligned}$$

Differentiating the integrand with respect to $f(q)$ and setting it equal to zero, we find the pair

$$f^*(q) = \sqrt{\frac{\varphi(q)}{\Gamma^*}} \quad (7)$$

and

$$\Gamma^* = \left(\frac{1}{2mvr} \int_{\mathcal{Q}} \sqrt{\varphi(q)} dq \right)^2$$

satisfy the Kuhn-Tucker necessary conditions for optimality [42], and since it is a convex program, these conditions are sufficient to insure global optimality. Upon substitution, (6) is proved. On the other hand, the constraint of unbiased service requires that

$$\mathbb{E}[T_j^{\text{detect}}|q_j = q] = \frac{1}{2f(q)} = \bar{T}^{\text{detect}}, \quad \forall q \in \mathcal{Q}.$$

Substituting into $\int_{\mathcal{Q}} f(q) dq \leq 2mvr$, we get

$$\frac{A}{2\bar{T}^{\text{detect}}} \leq 2mvr.$$

Rearranging and recalling $\bar{T}_{\text{opt}} \geq \bar{T}^{\text{detect}}$, we arrive at (5). ■

Oftentimes, a tight lower bound offers insight into the optimal solution of a problem. Assuming that this lower bound is tight, Eq. (7) suggests that in the spatially biased small sensing-range case, the optimal policy searches a point q at regular intervals at a relative frequency proportional to $\sqrt{\varphi(q)}$.

Moreover, we have presented new lower bounds on the optimal system time related to the searching capability of the agents and the necessary structure of any stabilizing policy. In the following sections, we will use these bounds to evaluate the performance of our proposed policies for the case of small sensor range.

IV. ALGORITHMS AND POLICIES FOR A SINGLE AGENT

In this section, we present four policies for the single-vehicle Limited Sensor DTRP, and prove their respective optimality in different limiting cases and classes of the problem, namely, the case of small sensor range (spatially unbiased and biased), and the case of heavy load (spatially unbiased and biased). To begin, we present two algorithms (subroutines used by the policies) by which an agent can service targets in a given region of the environment. The first is designed for the small-sensor case, and the second is designed for the heavy load case. We analyze their properties in their respective cases.

In the following, we consider a convex subregion $\mathcal{S} \subseteq \mathcal{Q}$ of area $A_{\mathcal{S}}$. The targets in \mathcal{S} are generated by a local Poisson process with time-intensity $\lambda_{\mathcal{S}} = \lambda\varphi(\mathcal{S})$ and spatial distribution $\varphi_{\mathcal{S}}(q) = \varphi(q)/\varphi(\mathcal{S})$ for all $q \in \mathcal{S}$. Note that $\varphi_{\mathcal{S}}$ is normalized such that $\varphi_{\mathcal{S}}(\mathcal{S}) = 1$. The first algorithm is designed for the case of small sensing-range.

SWEEP-SERVICE

The description of this algorithm requires the use of an inertial Cartesian coordinate frame. The algorithm is defined as follows.

- Partition \mathcal{S} into elements of width $2r$ with lines parallel to the x -axis.
- Define a strip as the bounding rectangle of an element of the partition, with sides parallel to the coordinate axes, and minimum side-lengths.
- Plan a path running along the longitudinal bisector of each strip, visiting all strips from top-to-bottom, connecting adjacent strip bisectors by their endpoints.
- Execute this path and visit targets as they are detected in the following manner. If a target is detected in the current strip and it is in front of the agent (with respect to the direction the agent is moving on the path), then the agent continues on the path until its position has the same x -coordinate as the target's. It then departs from the path, moving directly to the target, returning to the point of departure, and continuing on the path. If a target is detected outside the current strip, or behind the agent, then it is ignored.

We now analyze the length $L_{\text{swp}}(\mathcal{S})$ of the path planned by the algorithm.

Proposition 3: The length $L_{\text{swp}}(\mathcal{S})$ of the path planned by SWEEP-SERVICE for region \mathcal{S} satisfies

$$\lim_{r \rightarrow 0^+} L_{\text{swp}}(\mathcal{S})r \leq \frac{A_{\mathcal{S}}}{2}$$

Proof: Consider a grid of squares with sides of length $2r$ parallel to the coordinate axes. Denote $N_{\text{sq}}(\mathcal{S})$ as the number of squares with nonzero intersection with \mathcal{S} . The sum of the

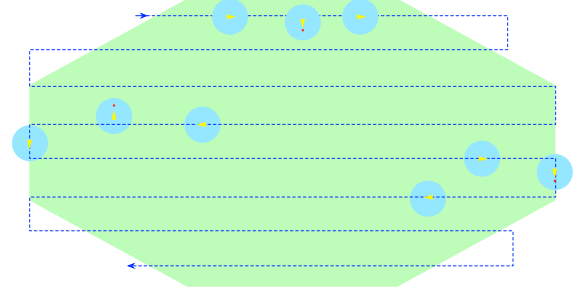


Fig. 1. Depiction of an agent executing SWEEP-SERVICE.

lengths of all strips can be bounded by $2rN_{\text{sq}}(\mathcal{S})$. Bound the region \mathcal{S} with a rectangle whose sides are parallel with the coordinate axes, and denote the length of its perimeter $P(\mathcal{S})$. The length of the path between the endpoints of the strip bisectors can be bounded above by $P(\mathcal{S})$. Thus, the total path length satisfies

$$L_{\text{swp}}(\mathcal{S}) \leq 2rN_{\text{sq}}(\mathcal{S}) + P(\mathcal{S}).$$

But the number of squares satisfies [43]

$$\lim_{r \rightarrow 0^+} N_{\text{sq}}(\mathcal{S})r^2 = \frac{A_{\mathcal{S}}}{4}$$

and so

$$\lim_{r \rightarrow 0^+} L_{\text{swp}}(\mathcal{S})r \leq \lim_{r \rightarrow 0^+} (2N_{\text{sq}}(\mathcal{S})r^2 + P(\mathcal{S})r) = \frac{A_{\mathcal{S}}}{2}. \quad \blacksquare$$

SNAPSHOT-TSP

The second algorithm is designed for the heavy load case. This algorithm requires that the subregion $\mathcal{S} \subseteq \mathcal{Q}$ has a size and shape such that it can be contained in the sensor footprint of the agent, i.e., there exists a position p such that $\|p - q\| \leq r$ for all $q \in \mathcal{S}$. Let p_{snap} be one such position. The algorithm is defined as follows.

- Move to location p_{snap} and take a snapshot, i.e., store in memory, the locations of all targets outstanding at the current time, called t_{snap} .
- Compute a minimum-length tour of all points in the snapshot.
- Generate a uniformly distributed (in terms of path length) random starting position on the tour (not necessarily a target's location).
- Randomly select a direction (clockwise or counterclockwise) with equal probability.
- Move to the starting position and execute the tour in the chosen direction, ignoring all targets that appear after t_{snap} .

We now analyze the length of the tour computed by the algorithm. Define the set of targets in the snapshot as $\mathcal{D}_{\text{snap}}$ and denote the cardinality of this set as n_{snap} .

Proposition 4: Assuming that all targets generated before some past time t_{clear} were cleared from \mathcal{S} , and the set of targets outstanding at the current time t is the set generated by the local Poisson process in the time interval $(t_{\text{clear}}, t]$ of

length $\Delta t = t - t_{\text{clear}}$, the length of the tour computed by SNAPSHOT-TSP satisfies

$$\lim_{\lambda \rightarrow \infty} \frac{\mathbb{E}[\text{ETSP}(\mathcal{D}_{\text{snap}})]}{\sqrt{\lambda}} = \beta \sqrt{\Delta t} \int_{\mathcal{S}} \sqrt{\varphi(q)} dq.$$

Proof: The points in the set $\mathcal{D}_{\text{snap}}$ are i.i.d. with compact support $\mathcal{S} \subset \mathbb{R}^2$, and $n_{\text{snap}} \rightarrow \infty$ almost surely as $\lambda \rightarrow \infty$. By Theorem 1

$$\lim_{n_{\text{snap}} \rightarrow \infty} \frac{\text{ETSP}(\mathcal{D}_{\text{snap}})}{\sqrt{n_{\text{snap}}}} = \beta \int_{\mathcal{S}} \sqrt{\varphi_{\mathcal{S}}(q)} dq.$$

Thus as $n_{\text{snap}} \rightarrow \infty$,

$$\begin{aligned} \mathbb{E}[\text{ETSP}(\mathcal{D}_{\text{snap}})] &= \mathbb{E} \left[\beta \sqrt{n_{\text{snap}}} \int_{\mathcal{S}} \sqrt{\varphi_{\mathcal{S}}(q)} dq \right] \\ &= \beta \mathbb{E}[\sqrt{n_{\text{snap}}}] \int_{\mathcal{S}} \sqrt{\varphi_{\mathcal{S}}(q)} dq. \end{aligned}$$

Define the length of the time interval over which $\mathcal{D}_{\text{snap}}$ was generated as $\Delta t = t_{\text{snap}} - t_{\text{clear}}$. The random variable n_{snap} is Poisson with mean $\mathbb{E}[n_{\text{snap}}] = \lambda_{\mathcal{S}} \Delta t$. By Jensen's inequality,

$$\mathbb{E}[\sqrt{n_{\text{snap}}}] \leq \sqrt{\mathbb{E}[n_{\text{snap}}]}.$$

However, as $\lambda_{\mathcal{S}} \rightarrow \infty$, the above inequality approaches equality, i.e.,

$$\lim_{\lambda_{\mathcal{S}} \rightarrow \infty} \frac{\mathbb{E}[\sqrt{n_{\text{snap}}}]}{\sqrt{\lambda_{\mathcal{S}}}} = \sqrt{\Delta t}$$

and thus

$$\lim_{\lambda_{\mathcal{S}} \rightarrow \infty} \frac{\mathbb{E}[\text{ETSP}(\mathcal{D}_{\text{snap}})]}{\sqrt{\lambda_{\mathcal{S}}}} = \beta \sqrt{\Delta t} \int_{\mathcal{S}} \sqrt{\varphi_{\mathcal{S}}(q)} dq.$$

Substituting $\lambda_{\mathcal{S}} = \lambda \varphi(\mathcal{S})$ and $\varphi_{\mathcal{S}}(q) = \varphi(q)/\varphi(\mathcal{S})$, we arrive at the claim. ■

The reader might find the choice of taking a snapshot at a particular instant and ignoring all targets generated thereafter peculiar. One might suggest that the agent could easily service newly generated targets whose locations happen to coincide with the vicinity of targets already in the current snapshot. However, with the described method, the sets of points in the snapshot are i.i.d. from the given Poisson process, and this allows us to apply Theorem 1 to the tour computed for each snapshot.

We now present four policies, each of which is designed for one of the four cases: small sensing-range (spatially biased and unbiased) and heavy load (spatially biased and unbiased). Some of the policies we present require a partition of the environment \mathcal{Q} into tiles $(\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_K)$, i.e., $\cup_{k=1}^K \mathcal{S}_k = \mathcal{Q}$ and $\mathcal{S}_k \cap \mathcal{S}_{\ell} = \emptyset$ if $k \neq \ell$. Each policy has different required properties of the partition, however, let us define some of the properties here. An equitable partition with respect to a measure $\psi : \mathcal{Q} \rightarrow \mathbb{R}_+$ is a partition such that $\psi(\mathcal{S}_k) = \psi(\mathcal{S}_{\ell})$ for all $k, \ell \in I_K$. Note that this condition implies $\psi(\mathcal{S}_k) = \psi(\mathcal{Q})/K$ for all $k \in I_K$. A convex partition is one whose subsets are convex.

A. The Unbiased Region Sweep (URS) Policy

The policy is defined in Algorithm 1. The index i is a label for the current phase of the policy.

for $i \leftarrow 1$ **to** ∞ **do**
 | Execute SWEEP-SERVICE on the environment \mathcal{Q}
end

Algorithm 1: URS Policy

Theorem 5: Let \bar{T}_{opt} be the optimal system time for the single-agent Limited Sensing DTRP over the class of spatially unbiased policies. Then the system time of a single agent operating on \mathcal{Q} under the URS policy satisfies

$$\frac{\bar{T}_{\text{URS}}}{\bar{T}_{\text{opt}}} = 1 \quad \text{as } r \rightarrow 0^+. \quad (8)$$

Proof: Define a phase of this policy as the time interval over which the agent performs one execution of SWEEP-SERVICE on the region \mathcal{Q} . Denote the length of the i -th phase by T_i^{phase} , and the number of targets visited during the i -th phase by n_i . Assuming that the policy is stabilizing, n_i is finite. The expected length of the i -th phase T_i^{phase} , conditioned on n_i , satisfies

$$\mathbb{E}[T_i^{\text{phase}} | n_i] \leq \frac{L_{\text{swp}}(\mathcal{Q}) + 2rn_i + \text{diam}(\mathcal{Q})}{v}.$$

Applying Proposition 3,

$$\begin{aligned} \lim_{r \rightarrow 0^+} \mathbb{E}[T_i^{\text{phase}} | n_i] r &\leq \\ \lim_{r \rightarrow 0^+} \frac{L_{\text{swp}}(\mathcal{Q})r + 2n_i r^2 + \text{diam}(\mathcal{Q})r}{v} &= \frac{A}{2v}. \end{aligned}$$

In other words, in the limit as $r \rightarrow 0^+$, the length of the phase does not depend on the number of targets serviced, so long as it is finite. In expectation, a target waits one half of a phase to be serviced, independent of its location. Therefore

$$\lim_{r \rightarrow 0^+} \bar{T}_{\text{URS}} r \leq \frac{A}{4v},$$

and moreover, the policy is spatially unbiased. Combining the above result with the lower bound on the optimal system time within the class of spatially unbiased policies in Theorem 2 and substituting $m = 1$, the claim is proved. ■

Theorem 5 shows that the optimal spatially unbiased policy for small sensing range simply searches the entire environment with equal frequency. In this case, the constraint on spatial bias does not allow the agent to leverage non-uniformity in the distribution of targets, φ , in order to lower the system time. We performed numerical experiments of the URS policy and results are shown in comparison with the lower bound (Eq. (5)) in Fig. IV-A on a log-log plot. The URS policy provides near-optimal performance for very small values of sensing radius r .

B. The Biased Tile Sweep (BTS) Policy

This policy requires that φ be a piecewise uniform density, i.e., $\mathcal{Q}_1, \mathcal{Q}_2, \dots, \mathcal{Q}_J$ be a partition of \mathcal{Q} such that $\varphi(q) = \mu_j \forall q \in \mathcal{Q}_j, j = 1, 2, \dots, J$. Let us assume that each subset \mathcal{Q}_j is convex, as a non-convex subset can be further partitioned

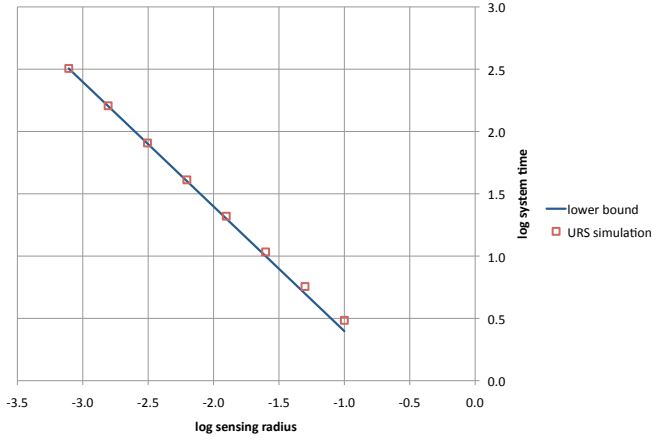


Fig. 2. Simulation performance of the URS policy for a single agent with unit velocity in a unit square environment and uniform spatial distribution. Results are compared with the theoretical lower bound (Eq. (5)) as radius r approaches zero on a log-log plot. The URS policy provides near-optimal performance for very small values of r .

into convex subsets. Let us denote $A_j = \text{Area}(\mathcal{Q}_j)$, $j = 1, 2, \dots, J$.

This policy requires a tiling of the environment \mathcal{Q} with the following properties. For some positive integer $K \in \mathbb{N}$, partition each subset \mathcal{Q}_j into $K_j = K/\sqrt{\mu_j}$ convex tiles, each of area $A_j/K_j = A_j\sqrt{\mu_j}/K$. We assume K is chosen large enough that an integer K_j can be found such that K/K_j is sufficiently close to $\sqrt{\mu_j}$. In other words, it requires a convex and equitable partition of each region \mathcal{Q}_j with respect to a constant measure ψ . This can be done with the following simple method. Partition \mathcal{Q}_j into strip-like tiles of equal measure with K_j parallel lines. Let us give the tiles of \mathcal{Q}_j an ordered labeling $\mathcal{S}_{j,1}, \mathcal{S}_{j,2}, \dots, \mathcal{S}_{j,K_j}$. The BTS policy is defined in Algorithm 2. The index i is a label for the current phase of the policy.

```

Initialize  $k_j \leftarrow 1$  for  $j = 1, 2, \dots, J$ 
for  $i \leftarrow 1$  to  $\infty$  do
  for  $j \leftarrow 1$  to  $J$  do
    Execute SWEEP-SERVICE on tile  $\mathcal{S}_{j,k_j}$ 
    if  $k_j < K_j$  then  $k_j \leftarrow (k_j + 1)$ 
    else  $k_j \leftarrow 1$ 
  end
end

```

Algorithm 2: BTS Policy

Example 6: In order to illustrate the application of the BTS policy on a specific problem instance, we consider the environment and piece-wise constant density function shown in Fig. 6. This example is made up of four subregions, \mathcal{Q}_j , each of constant density, μ_j , indicated in the drawing. We do not specify the areas of the subregions as their magnitudes are not relevant to the algorithm. For the example to be well posed, we can assume that the areas are such that $\varphi(\mathcal{Q}) = 1$. In fact, the absolute values of the μ_j 's are not relevant either. In order to apply the BTS policy, all we need to know are the relative magnitudes of the μ_j 's. In other words, how much

more likely is it that a target appears in \mathcal{Q}_1 rather than \mathcal{Q}_2 ? Towards the end of making the relative magnitudes clear, this example's lowest density subregion has a density of 1. Any given piecewise-constant density can be normalized to such a form.

The first step of the algorithm is to choose our scaling constant, a positive integer $K \in \mathbb{N}$. Since the highest density is $\mu_1 = 36$ and $K_1 = K/\sqrt{\mu_1}$ in order to ensure that $K_1 \geq 1$, we must set $K \geq 6$. The beneficial aspect of setting K arbitrarily large is that it allows us to make K/K_j arbitrarily close to $\sqrt{\mu_j}$ for each j . But in practice, there is a cost associated with increasing K : it increases the frequency of transition moves between sweeping tiles. In the limit as $r \rightarrow 0^+$ these transition costs are negligible as the time required to sweep an individual tile dwarfs them, but for small but fixed r , these transition costs must be balanced with the benefit of designing the ideal ratio's between the K_j 's. In this particular example, we don't have to make a trade-off between the transition costs and the ideal ratio's of the K_j 's. We can set K to its minimum feasible value, $K = 6$, and then $K_j = 6/\sqrt{\mu_j}$ for each j , resulting in

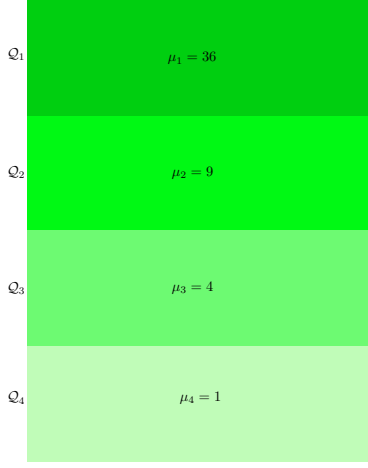
$$\begin{aligned}
 K_1 &= 1, \\
 K_2 &= 2, \\
 K_3 &= 3, \\
 K_4 &= 6.
 \end{aligned}$$

Each subregion \mathcal{Q}_j is divided into K_j tiles of equal measure, $\mathcal{S}_{j,1}, \mathcal{S}_{j,2}, \dots, \mathcal{S}_{j,K_j}$, as shown in Fig. 6. Then the agent performs SWEEP-SERVICE on the tiles in the following repeating sequence. During each phase, one tile from each subregion is swept:

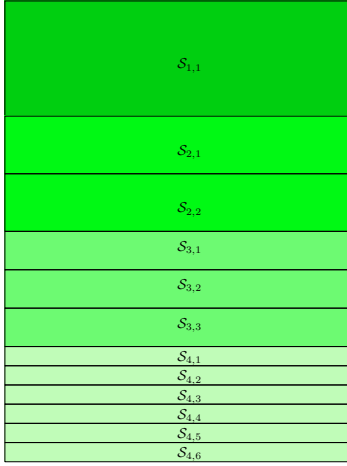
$$\begin{aligned}
 \text{phase}_1 &: \{\mathcal{S}_{1,1}, \mathcal{S}_{2,1}, \mathcal{S}_{3,1}, \mathcal{S}_{4,1}\}, \\
 \text{phase}_2 &: \{\mathcal{S}_{1,1}, \mathcal{S}_{2,2}, \mathcal{S}_{3,2}, \mathcal{S}_{4,2}\}, \\
 \text{phase}_3 &: \{\mathcal{S}_{1,1}, \mathcal{S}_{2,1}, \mathcal{S}_{3,3}, \mathcal{S}_{4,3}\}, \\
 \text{phase}_4 &: \{\mathcal{S}_{1,1}, \mathcal{S}_{2,2}, \mathcal{S}_{3,1}, \mathcal{S}_{4,4}\}, \\
 \text{phase}_5 &: \{\mathcal{S}_{1,1}, \mathcal{S}_{2,1}, \mathcal{S}_{3,2}, \mathcal{S}_{4,5}\}, \\
 \text{phase}_6 &: \{\mathcal{S}_{1,1}, \mathcal{S}_{2,2}, \mathcal{S}_{3,3}, \mathcal{S}_{4,6}\}.
 \end{aligned}$$

Again, the relative magnitudes of the areas of the subregions are irrelevant to the algorithm. The subregions in this example are of equal area only for clarity and simplicity. Of course, the absolute magnitudes of the areas relative to the agent's sensing radius will directly influence the system time achieved by the algorithm. Also, the subregions need not be convex as convexity is not required to compute an efficient path coverage. Furthermore, the subregions need not be connected. We only require that the time required to travel between them be negligible relative to the time required to perform path coverage sweeps on them.

Theorem 7: If φ is a piecewise uniform density and \bar{T}_{opt} is the optimal system time for the single-agent Limited Sensing DTRP over the class of spatially biased policies, then the system time of a single agent operating on \mathcal{Q} under the BTS



(a) Example environment and distribution.



(b) Tiling created by BTS policy.

Fig. 3. Example environment and resultant tiling created by BTS policy.

policy satisfies

$$\frac{\bar{T}_{\text{BTS}}}{T_{\text{opt}}} = 1 \quad \text{as } r \rightarrow 0^+. \quad (9)$$

Proof: The total distance traveled between tiles during a phase is no more than $J \text{diam}(\mathcal{Q})$. Assuming that the policy is stabilizing, the number of targets serviced during the i -th phase n_i is finite. The expected length of the i -th phase T_i^{phase} , conditioned on n_i , satisfies

$$\mathbb{E} [T_i^{\text{phase}} | n_i] \leq \frac{\sum_{j=1}^J L_{\text{swp}}(\mathcal{S}_{j,k_j}) + 2rn_i + J \text{diam}(\mathcal{Q})}{v}.$$

Applying Proposition 3,

$$\begin{aligned} \lim_{r \rightarrow 0^+} \mathbb{E} [T_i^{\text{phase}} | n_i] r &\leq \\ \lim_{r \rightarrow 0^+} \frac{\sum_{j=1}^J L_{\text{swp}}(\mathcal{S}_{j,k_j}) r}{v} + \lim_{r \rightarrow 0^+} \frac{2n_i r^2 + J \text{diam}(\mathcal{Q}) r}{v} &= \frac{\sum_{j=1}^J A_j / K_j}{2v} \\ &= \frac{\sum_{j=1}^J A_j \sqrt{\mu_j}}{2Kv}. \end{aligned}$$

Conditioned upon its location $q \in \mathcal{Q}_j$, a target waits one half of K_j phases to be serviced,

$$\begin{aligned} \mathbb{E} [\bar{T}_{\text{BTS}} | q \in \mathcal{Q}_j] &= \frac{1}{2} K_j T^{\text{phase}} \\ &= \frac{1}{2} \frac{K}{\sqrt{\mu_j}} T^{\text{phase}}. \end{aligned}$$

Noting that $\Pr[q \in \mathcal{Q}_j] = \mu_j A_j$ and unconditioning on $q \in \mathcal{Q}_j$ to find the system time,

$$\begin{aligned} \bar{T}_{\text{BTS}} &= \sum_{j=1}^J \Pr[q \in \mathcal{Q}_j] \cdot \mathbb{E} [\bar{T}_{\text{BTS}} | q \in \mathcal{Q}_j] \\ &= \sum_{j=1}^J (A_j \mu_j) \cdot \left(\frac{1}{2} \frac{K}{\sqrt{\mu_j}} T^{\text{phase}} \right) \\ &= \frac{KT^{\text{phase}}}{2} \sum_{j=1}^J A_j \sqrt{\mu_j}. \end{aligned}$$

Thus,

$$\begin{aligned} \lim_{r \rightarrow 0^+} \bar{T}_{\text{BTS}} r &= \left(\frac{K}{2} \sum_{j=1}^J A_j \sqrt{\mu_j} \right) \lim_{r \rightarrow 0^+} T^{\text{phase}} r \\ &= \frac{1}{4v} \left(\sum_{j=1}^J A_j \sqrt{\mu_j} \right)^2. \end{aligned}$$

If φ is a piecewise uniform density and $m = 1$, the lower bound on the optimal system time within the class of spatially biased policies in Theorem 2 takes on the form

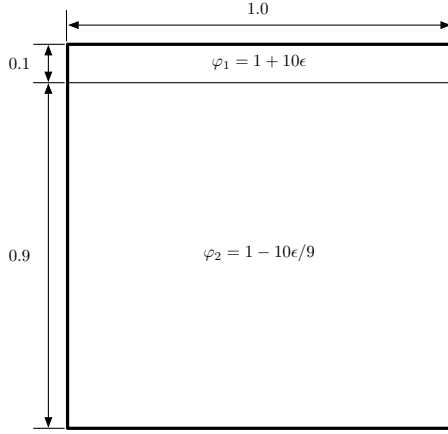
$$\lim_{r \rightarrow 0^+} \bar{T}_{\text{opt}} r \geq \frac{1}{4v} \left(\sum_{j=1}^J A_j \sqrt{\mu_j} \right)^2.$$

Combining, the claim is proved. \blacksquare

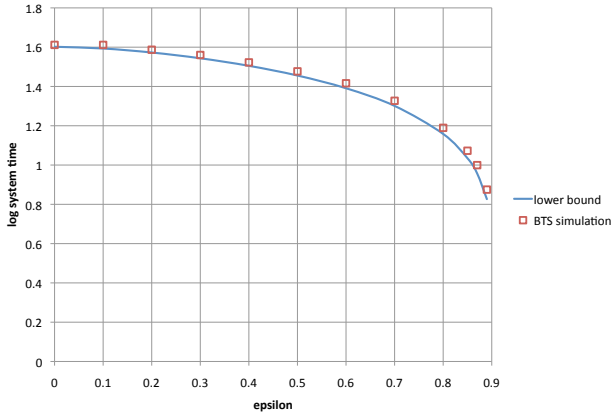
Theorem 7 shows that the optimal spatially biased policy for small sensor-range searches a point in the environment at regular intervals at a relative frequency proportional to $\sqrt{\varphi(q)}$, as was suggested by Eq. (7) in the proof of the corresponding spatially biased lower bound in Theorem 2.

We performed numerical experiments of the BTS policy for a single agent. As shown in Fig. IV-B (a), the spatial distribution of the target-generation process was piece-wise uniform with a density of $\varphi_1 = 1 + 10\epsilon$ in the smaller region of area 0.1 and $\varphi_2 = 1 + 10\epsilon/9$ in the larger region of area 0.9. We varied ϵ from 0 to 0.89. This tested the algorithm's

performance under a large range of spatial distributions: from uniform ($\epsilon = 0$) to one in which targets appear in the smaller region with a 99% probability ($\epsilon = 0.89$). Results are shown in comparison with the lower bound (Eq. 6) in Fig. IV-B (b) on a semi-log plot. The BTS policy provides near-optimal performance for a large range of spatial distributions, i.e., the BTS policy adapts the distribution of the searching agent's position in order to exploit the spatially biased target-generation process and thereby reduce the expected system time overall. In other words, the searching agent provides higher quality of service to the targets in the higher density regions. Specifically, under the BTS policy, a target's expected quality of service will scale (relative to other targets in other regions) with the inverse square root of its region's density.



(a) Drawing of the environment and spatial distribution (parameterized using ϵ) used for the numerical experiments on the BTS policy.



(b) Simulation performance of the BTS policy for a single agent in a unit square environment with unit velocity and sensing radius $r = 0.00625$. Results are compared with the theoretical lower bound (Eq. 6) for varying spatial distribution on a semi-log plot. The parameter ϵ was varied from 0 to 0.89, i.e., the spatial distribution varied from uniform to one in which 99% of the incidents occurring in the subregion with 10% of the area.

Fig. 4. Numerical experiments of the BTS policy.

C. The Unbiased Tile TSP (UTTSP) Policy

This policy requires a tiling of the environment \mathcal{Q} with the following properties. For some positive integer $K \in \mathbb{N}$,

partition \mathcal{Q} into tiles $\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_K$ such that

$$\int_{\mathcal{S}_k} \sqrt{\varphi(q)} dq = \frac{1}{K} \int_{\mathcal{Q}} \sqrt{\varphi(q)} dq, \quad k = 1, 2, \dots, K.$$

In other words, it requires a convex and equitable partition of the environment \mathcal{Q} with respect to the measure ψ where $\psi(q) = \sqrt{\varphi(q)}$ for all $q \in \mathcal{Q}$. Furthermore, the size and shape of each tile must be such that it can be contained in the sensor footprint of an agent, i.e., for each \mathcal{S}_k there exists a point p_k such that $\|p_k - q\| \leq r$ for all $q \in \mathcal{S}_k$. For example, if each tile can be bounded by a rectangle, neither of whose side-lengths exceeds $r/\sqrt{2}$, then this property is achieved. A tiling with all these properties can be constructed with the following simple method. First, partition \mathcal{Q} into strip-like tiles of equal measure with K_1 parallel lines. If K_1 is sufficiently large, then the width of the thinnest tile is less than or equal to $r/\sqrt{2}$. Next, partition each of the K_1 strip-like tiles into K_2 tiles of equal measure using lines perpendicular to the first set. If K_2 is sufficiently large, then the height of all tiles is less than or equal to $r/\sqrt{2}$. The UTTSP policy is defined in Algorithm 3. The index i is a label for the current phase of the policy.

```

for  $i \leftarrow 1$  to  $\infty$  do
  for  $k \leftarrow 1$  to  $K$  do
    Execute SNAPSHOT-TSP on tile  $\mathcal{S}_k$ 
  end
end

```

Algorithm 3: UTTSP Policy

Theorem 8: Let \bar{T}_{opt} be the optimal system time for the single-agent Limited Sensing DTRP over the class of spatially unbiased policies. Then the system time of a single agent operating on \mathcal{Q} under the UTTSP policy satisfies

$$\frac{\bar{T}_{\text{UTTSP}}}{\bar{T}_{\text{opt}}} = 1 \quad \text{as } \lambda \rightarrow \infty. \quad (10)$$

Proof: Let us denote $T_{k,i}^{\text{tsp}}$ as the time required to execute the tour of the targets in tile \mathcal{S}_k computed by SNAPSHOT-TSP in the i -th phase. Including the time traveling between tiles, we note that

$$T_i^{\text{phase}} \leq \sum_{k=1}^K T_{k,i}^{\text{tsp}} + \frac{K \text{diam}(\mathcal{Q})}{v}.$$

Applying Proposition 4, the time required for the first tile in the $(i+1)$ -th phase $T_{1,i+1}^{\text{tsp}}$, conditioned on the length of the i -th phase T_i^{phase} , satisfies

$$\begin{aligned} \lim_{\lambda \rightarrow \infty} \frac{\mathbb{E}[T_{1,i+1}^{\text{tsp}} | T_i^{\text{phase}}]}{\sqrt{\lambda}} &= \frac{\beta \sqrt{T_i^{\text{phase}}}}{v} \int_{\mathcal{S}_1} \sqrt{\varphi(q)} dq \\ &= \frac{\beta \sqrt{T_i^{\text{phase}}}}{Kv} \int_{\mathcal{Q}} \sqrt{\varphi(q)} dq. \end{aligned}$$

Since the choice of the first tile, and the epoch of the phase, is arbitrary, the expected time to visit all targets in a tile is

uniform over all tiles. Summing over all tiles, we get

$$\begin{aligned} \lim_{\lambda \rightarrow \infty} \frac{\mathbb{E} [T_{i+1}^{\text{phase}} | T_i^{\text{phase}}]}{\sqrt{\lambda}} &= \\ K \cdot \lim_{\lambda \rightarrow \infty} \frac{\mathbb{E} [T_{1,i+1}^{\text{tsp}} | T_i^{\text{phase}}]}{\sqrt{\lambda}} + \lim_{\lambda \rightarrow \infty} \frac{K \text{diam}(\mathcal{Q})}{v\sqrt{\lambda}} &= \\ = \frac{\beta\sqrt{T_i^{\text{phase}}}}{v} \int_{\mathcal{Q}} \sqrt{\varphi(q)} dq. \end{aligned}$$

From the equation above, it can be verified that

$$\begin{aligned} \mathbb{E} [T_{i+1}^{\text{phase}} | T_i^{\text{phase}}] &\geq T_i^{\text{phase}} \\ \text{if } T_i^{\text{phase}} &\leq \frac{\beta^2 \lambda}{v^2} \left(\int_{\mathcal{Q}} \sqrt{\varphi(q)} dq \right)^2 \end{aligned}$$

and

$$\begin{aligned} \mathbb{E} [T_{i+1}^{\text{phase}} | T_i^{\text{phase}}] &\leq T_i^{\text{phase}} \\ \text{if } T_i^{\text{phase}} &\geq \frac{\beta^2 \lambda}{v^2} \left(\int_{\mathcal{Q}} \sqrt{\varphi(q)} dq \right)^2. \end{aligned}$$

Thus, the sequence of phases exhibits a fixed-point, from which

$$\lim_{i \rightarrow \infty} \mathbb{E} [T_i^{\text{phase}}] = \frac{\beta^2 \lambda}{v^2} \left(\int_{\mathcal{Q}} \sqrt{\varphi(q)} dq \right)^2.$$

Substituting, as $i \rightarrow \infty$,

$$\begin{aligned} \mathbb{E} [T_{1,i}^{\text{tsp}}] &= \frac{\beta\sqrt{\lambda}\sqrt{T_i^{\text{phase}}}}{Kv} \int_{\mathcal{Q}} \sqrt{\varphi(q)} dq \\ &= \frac{\beta^2 \lambda}{Kv^2} \left(\int_{\mathcal{Q}} \sqrt{\varphi(q)} dq \right)^2 \\ &= \frac{T_i^{\text{phase}}}{K}. \end{aligned}$$

Note that the above is quantity is independent of the specific tile. In expectation, a target waits one half of a phase to enter a snapshot. Because of the randomization in starting point and direction performed in SNAPSHOT-TSP, in expectation, a target waits one half of the time required to visit all targets in its snapshot. Thus,

$$\begin{aligned} \mathbb{E} [\bar{T}_{\text{UTTSP}}] &= \frac{1}{2} T^{\text{phase}} + \frac{1}{2} \frac{T^{\text{phase}}}{K} \\ &= \left(1 + \frac{1}{K} \right) \frac{\beta^2 \lambda}{2v^2} \left(\int_{\mathcal{Q}} \sqrt{\varphi(q)} dq \right)^2. \end{aligned}$$

Since the system time is independent of the location of the target, the policy is spatially unbiased. Using the optimal system time for the full-information DTRP in heavy load within the class of spatially unbiased policies in Eq. (3) with $m = 1$ as a lower bound, we see that for large K , the claim is proved. ■

Although the constraint on spatial bias does not allow the policy to service denser regions with higher frequency, Theorem 8 shows that non-uniformity in the spatial distribution of targets, φ , still leads to a lowering of the optimal system time. This is due to the efficiency gained by the ETSP tours due to

non-uniformity, evident in Theorem 1.

D. The Biased Tile TSP (BTTSP) Policy

This policy requires that φ be a piecewise uniform density, i.e., $\mathcal{Q}_1, \mathcal{Q}_2, \dots, \mathcal{Q}_J$ be a partition of \mathcal{Q} such that $\varphi(q) = \mu_j \forall q \in \mathcal{Q}_j, j = 1, 2, \dots, J$. Let us assume that each subset \mathcal{Q}_j is convex, as a non-convex subset can be further partitioned into convex subsets. Let us denote $A_j = \text{Area}(\mathcal{Q}_j)$, $j = 1, 2, \dots, J$.

This policy requires a tiling of the environment \mathcal{Q} with the following properties. For some positive integer $K \in \mathbb{N}$, partition each subset \mathcal{Q}_j into $K_j = K/\mu_j^{1/3}$ convex tiles, each of area $A_j/K_j = A_j \mu_j^{1/3}/K$. We assume K is chosen large enough that an integer K_j can be found such that K/K_j is sufficiently close to $\mu_j^{1/3}$. Furthermore, the size and shape of each tile must be such that it can be contained in the sensor footprint of an agent, i.e., for each \mathcal{S}_k there exists a point p_k such that $\|p_k - q\| \leq r$ for all $q \in \mathcal{S}_k$. The grid-like equitable tiling described for the UTTSP policy, applied here, would require that each K_j be factorable into possibly large numbers K_1 and K_2 . This is undesirable because the numbers K_j must maintain specific ratios related to the density μ_j in their domain \mathcal{Q}_j as described above. However, this scenario does have one simpler facet: the density functions within each \mathcal{Q}_j are constant. One example of a method for reaching convex and equitable partitions is given in [44]. Heuristically speaking, the proposed algorithms converge to configurations in which all cells are approximately hexagonal for constant measure ψ and large K_j . Hence, for sufficiently large K , these hexagonal tiles fit within a circle of radius r . Let us give the tiles of \mathcal{Q}_j an ordered labeling $\mathcal{S}_{j,1}, \mathcal{S}_{j,2}, \dots, \mathcal{S}_{j,K_j}$. The BTTSP policy is defined in Algorithm 4. The index i is a label for the current phase of the policy.

```

Initialize  $k_j \leftarrow 1$  for  $j = 1, 2, \dots, J$ 
for  $i \leftarrow 1$  to  $\infty$  do
  for  $j \leftarrow 1$  to  $J$  do
    Execute SNAPSHOT-TSP on tile  $\mathcal{S}_{j,k_j}$ 
    if  $k_j < K_j$  then  $k_j \leftarrow (k_j + 1)$ 
    else  $k_j \leftarrow 1$ 
  end
end

```

Algorithm 4: BTTSP Policy

Theorem 9: If φ is a piecewise uniform density and \bar{T}_{opt} is the optimal system time for the single-agent Limited Sensing DTRP over the class of spatially biased policies, then the system time of a single agent operating on \mathcal{Q} under the BTTSP policy satisfies

$$\frac{\bar{T}_{\text{BTTSP}}}{\bar{T}_{\text{opt}}} = 1 \quad \text{as } \lambda \rightarrow \infty. \quad (11)$$

Proof: Assume, for now, that the sequence of phases exhibits a fixed-point, from which

$$\lim_{i \rightarrow \infty} \mathbb{E} [T_i^{\text{phase}}] = T_{\text{ss}}^{\text{phase}},$$

where T_{ss}^{phase} denotes the steady-state phase length. Denote $T_{j,k_j,i}^{\text{tsp}}$ as the time required to execute the tour of the targets in tile \mathcal{S}_{j,k_j} computed by SNAPSHOT-TSP in the i -th phase. Any tile in \mathcal{Q}_j waits K_j phases between snapshots. Applying Proposition 4 with the fixed-point assumption, as $i \rightarrow \infty$,

$$\begin{aligned} \lim_{\lambda \rightarrow \infty} \frac{\mathbb{E}[T_{j,k_j,i}^{\text{tsp}}]}{\sqrt{\lambda}} &= \frac{\beta}{v} \sqrt{K_j T_{ss}^{\text{phase}}} \int_{\mathcal{S}_{j,k_j}} \sqrt{\varphi(q)} dq \\ &= \frac{\beta}{v} \sqrt{K_j T_{ss}^{\text{phase}}} \frac{A_j}{K_j} \sqrt{\mu_j} \\ &= \frac{\beta}{v} \sqrt{\frac{T_{ss}^{\text{phase}}}{K}} A_j \mu_j^{2/3}. \end{aligned}$$

In other words, the fixed-point assumption implies that

$$\lim_{i \rightarrow \infty} \mathbb{E}[T_{j,k_j,i}^{\text{tsp}}] = T_{j,k_j,ss}^{\text{tsp}} = \frac{\beta}{v} \sqrt{\frac{\lambda T_{ss}^{\text{phase}}}{K}} A_j \mu_j^{2/3},$$

for $k_j = 1, 2, \dots, K_j$. Summing over all tiles in a phase, and including distance traveled between tiles,

$$\begin{aligned} \lim_{\lambda \rightarrow \infty} \frac{\mathbb{E}[T_{ss}^{\text{phase}}]}{\sqrt{\lambda}} &= \sum_{j=1}^J \lim_{\lambda \rightarrow \infty} \frac{\mathbb{E}[T_{j,k_j,ss}^{\text{tsp}}]}{\sqrt{\lambda}} + \lim_{\lambda \rightarrow \infty} \frac{J \cdot \text{diam}(\mathcal{Q})}{v\sqrt{\lambda}} \\ &= \frac{\beta}{v} \sqrt{\frac{T_{ss}^{\text{phase}}}{K}} \sum_{j=1}^J A_j \mu_j^{2/3}. \end{aligned}$$

The above implies that

$$T_{ss}^{\text{phase}} = \frac{\beta^2 \lambda}{K v^2} \left(\sum_{j=1}^J A_j \mu_j^{2/3} \right)^2.$$

We now investigate the expected system time of a target, conditioned upon its location $q \in \mathcal{Q}_j$. The time a target waits before entering a snapshot, T_{snap}^- , is one half of K_j phases, i.e.,

$$\begin{aligned} \mathbb{E}[T_{\text{snap}}^- | q \in \mathcal{Q}_j] &= \frac{1}{2} K_j T_{ss}^{\text{phase}} \\ &= \frac{1}{2} K T_{ss}^{\text{phase}} \mu_j^{-1/3}. \end{aligned}$$

Unconditioning on the location of the target,

$$\begin{aligned} \mathbb{E}[T_{\text{snap}}^-] &= \sum_{j=1}^J \Pr[q \in \mathcal{Q}_j] \cdot \mathbb{E}[T_{\text{snap}}^- | q \in \mathcal{Q}_j] \\ &= \sum_{j=1}^J (A_j \mu_j) \cdot \left(\frac{1}{2} K T_{ss}^{\text{phase}} \mu_j^{-1/3} \right) \\ &= \frac{1}{2} K T_{ss}^{\text{phase}} \sum_{j=1}^J A_j \mu_j^{2/3} \\ &= \frac{\beta^2 \lambda}{2 v^2} \left(\sum_{j=1}^J A_j \mu_j^{2/3} \right)^3. \end{aligned}$$

Because of the randomization in starting point and direction performed in SNAPSHOT-TSP, in expectation, the time a target waits after entering a snapshot, T_{snap}^+ , is one half of

the time required to visit all targets in its snapshot.

$$\begin{aligned} \mathbb{E}[T_{\text{snap}}^+ | q \in \mathcal{Q}_j] &= \frac{1}{2} T_{j,k_j,ss}^{\text{tsp}} \\ &= \frac{1}{2} \frac{\beta}{v} \sqrt{\frac{\lambda T_{ss}^{\text{phase}}}{K}} A_j \mu_j^{2/3}. \end{aligned}$$

Unconditioning on the location of the target,

$$\begin{aligned} \mathbb{E}[T_{\text{snap}}^+] &= \sum_{j=1}^J \Pr[q \in \mathcal{Q}_j] \cdot \mathbb{E}[T_{\text{snap}}^+ | q \in \mathcal{Q}_j] \\ &= \sum_{j=1}^J (A_j \mu_j) \cdot \left(\frac{1}{2} \frac{\beta}{v} \sqrt{\frac{\lambda T_{ss}^{\text{phase}}}{K}} A_j \mu_j^{2/3} \right) \\ &= \frac{\beta}{2v} \sqrt{\frac{\lambda T_{ss}^{\text{phase}}}{K}} \sum_{j=1}^J A_j^2 \mu_j^{5/3} \\ &= \frac{\beta^2 \lambda}{2 v^2} \frac{1}{K} \left(\sum_{j=1}^J A_j \mu_j^{2/3} \right) \left(\sum_{j=1}^J A_j^2 \mu_j^{5/3} \right). \end{aligned}$$

Combining,

$$\begin{aligned} \bar{T}_{\text{BTTSP}} &= \mathbb{E}[T_{\text{snap}}^-] + \mathbb{E}[T_{\text{snap}}^+] \\ &= \frac{\beta^2 \lambda}{2 v^2} \left(\sum_{j=1}^J A_j \mu_j^{2/3} \right)^3 \\ &\quad + \frac{1}{K} \frac{\beta^2 \lambda}{2 v^2} \left(\sum_{j=1}^J A_j \mu_j^{2/3} \right) \left(\sum_{j=1}^J A_j^2 \mu_j^{5/3} \right). \end{aligned}$$

Namely, for large K , we have

$$\lim_{\lambda \rightarrow \infty} \frac{\bar{T}_{\text{BTTSP}}}{\lambda} = \frac{\beta^2}{2 v^2} \left(\sum_{j=1}^J A_j \mu_j^{2/3} \right)^3.$$

If φ is a piecewise uniform density and $m = 1$, then the optimal system time for the full-information DTRP in heavy load within the class of spatially biased policies in Eq. (4) takes on the form

$$\lim_{\lambda \rightarrow \infty} \frac{\bar{T}_{\text{opt}}}{\lambda} = \frac{\beta^2}{2 v^2} \left(\sum_{j=1}^J A_j \mu_j^{2/3} \right)^3. \quad (12)$$

Using the above as a lower bound for the Limited Sensor DTRP in heavy load, we see that for large K , the claim is proved. ■

Theorem 9 shows that the optimal spatially biased policy for heavy load services targets in the vicinity of a point q at regular intervals at a relative frequency proportional to $\varphi(q)^{1/3}$.

V. MULTIPLE AGENTS WITH EQUITABLE REGIONS OF DOMINANCE

We have presented four algorithms and proven them optimal in four respective cases of the single-vehicle Limited Sensing DTRP. We wish to adapt these policies to the multiple-vehicle scenario, retaining optimality, with minimal communication and collaboration among the agents. Towards this end, consider the following strategy. Given a single-vehicle policy

π , partition the environment \mathcal{Q} into regions of dominance $\mathcal{V}(\mathcal{Q}) = (\mathcal{V}_1, \mathcal{V}_2, \dots, \mathcal{V}_m)$ where $\cup_{\ell=1}^m \mathcal{V}_\ell = \mathcal{Q}$ and $\mathcal{V}_\ell \cap \mathcal{V}_h = \emptyset$ if $\ell \neq h$. Each vehicle executes policy π on its own region of dominance. In other words, agent ℓ is responsible for all targets appearing in \mathcal{V}_ℓ and ignores all others.

The control policies to reach convex and equitable partitions proposed in [44] are distributed and can be performed through m mobile agents, where each agent only need communicate with agents in neighboring cells of the partition. We do not pursue this facet of the multi-agent system further, as these methods are available and well studied. In the following theorems, we show that if the regions of dominance are a convex and equitable partition of \mathcal{Q} with respect to a designed measure ψ , then the decentralized strategy achieves an optimal system time. However, the appropriate measure ψ depends on the problem parameters (small sensing radius or heavy load) and spatial constraints (biased or unbiased). Moreover, it is a function of the spatial distribution of targets φ in three of the four cases.

Some of the single-vehicle policies in the previous section require the knowledge of certain properties of the environment such as its target generation rate, and its spatial target distribution. A single-vehicle policy operating on a subregion $\mathcal{V}_\ell \subseteq \mathcal{Q}$ takes as input the local target-generation process of \mathcal{V}_ℓ with time intensity $\lambda_\ell = \lambda\varphi(\mathcal{V}_\ell)$ and spatial distribution $\varphi_\ell : \mathcal{V}_\ell \rightarrow \mathbb{R}_+$ where $\varphi_\ell(q) = \varphi(q)/\varphi(\mathcal{V}_\ell)$ if $q \in \mathcal{V}_\ell$ and $\varphi_\ell(q) = 0$ otherwise. Note that φ_ℓ is normalized such that $\varphi_\ell(\mathcal{V}_\ell) = 1$.

Theorem 10: Let \bar{T}_{opt} be the optimal system time for the m -agent Limited Sensing DTRP over the class of spatially unbiased policies, and let $\mathcal{V}(\mathcal{Q})$ be a convex and equitable partition of \mathcal{Q} with respect to measure ψ where $\psi(q) = 1$ for all $q \in \mathcal{Q}$. If each agent $\ell \in I_m$ operates on its own region of dominance \mathcal{V}_ℓ under the URS policy, then the system time satisfies

$$\frac{\bar{T}_{\text{URS/ERD}}}{\bar{T}_{\text{opt}}} = 1 \quad \text{as } r \rightarrow 0^+. \quad (13)$$

Proof: Denoting $A_\ell = \text{Area}(\mathcal{V}_\ell)$, we apply Theorem 5 to find the system time of a target, conditioned upon its location,

$$\mathbb{E}[T_{\text{URS/ERD}}|q \in \mathcal{V}_\ell] = \frac{A_\ell}{4vr}.$$

Since $\psi(\mathcal{V}_\ell) = A_\ell$, the equitable partition implies that $A_\ell = A/m$ for all $\ell \in I_m$. Hence

$$\mathbb{E}[T_{\text{URS/ERD}}|q \in \mathcal{V}_\ell] = \frac{A}{4mvr},$$

and since the above is independent of the targets location, it is in fact the unconditioned system time. Moreover this implies that the strategy is spatially unbiased. Combining with Theorem 2, the claim is proved. ■

Theorem 11: Let φ be a piecewise uniform density, let \bar{T}_{opt} be the optimal system time for the m -agent Limited Sensing DTRP over the class of spatially biased policies, and let $\mathcal{V}(\mathcal{Q})$ be a convex and equitable partition of \mathcal{Q} with respect to measure ψ where $\psi(q) = \sqrt{\varphi(q)}$ for all $q \in \mathcal{Q}$. If each agent $\ell \in I_m$ operates on its own region of dominance \mathcal{V}_ℓ under the

BTS policy, then the system time satisfies

$$\frac{\bar{T}_{\text{BTS/ERD}}}{\bar{T}_{\text{opt}}} = 1 \quad \text{as } r \rightarrow 0^+. \quad (14)$$

Proof: We apply Theorem 7 to find the system time of a target, conditioned upon its location,

$$\mathbb{E}[T_{\text{BTS/ERD}}|q \in \mathcal{V}_\ell] = \frac{1}{4vr} \left(\int_{\mathcal{V}_\ell} \sqrt{\varphi_\ell(q)} dq \right)^2.$$

Substituting $\varphi_\ell(q) = \varphi(q)/\varphi(\mathcal{V}_\ell)$ we get

$$\mathbb{E}[T_{\text{BTS/ERD}}|q \in \mathcal{V}_\ell] = \frac{1}{4vr} \frac{1}{\varphi(\mathcal{V}_\ell)} \left(\int_{\mathcal{V}_\ell} \sqrt{\varphi(q)} dq \right)^2.$$

Unconditioning on the location of the target,

$$\begin{aligned} \bar{T}_{\text{BTS/ERD}} &= \sum_{\ell=1}^m \Pr[q \in \mathcal{V}_\ell] \cdot \mathbb{E}[T_{\text{BTS/ERD}}|q \in \mathcal{V}_\ell] \\ &= \sum_{\ell=1}^m \varphi(\mathcal{V}_\ell) \cdot \frac{1}{4vr} \frac{1}{\varphi(\mathcal{V}_\ell)} \left(\int_{\mathcal{V}_\ell} \sqrt{\varphi(q)} dq \right)^2 \\ &= \frac{1}{4vr} \sum_{\ell=1}^m \left(\int_{\mathcal{V}_\ell} \sqrt{\varphi(q)} dq \right)^2. \end{aligned}$$

But the equitable partition with respect to $\psi(q) = \sqrt{\varphi(q)}$ implies that

$$\int_{\mathcal{V}_\ell} \sqrt{\varphi(q)} dq = \frac{1}{m} \int_{\mathcal{Q}} \sqrt{\varphi(q)} dq, \quad \forall \ell \in I_m,$$

and so

$$\bar{T}_{\text{BTS/ERD}} = \frac{1}{4mvr} \left(\int_{\mathcal{Q}} \sqrt{\varphi(q)} dq \right)^2.$$

Combining with Theorem 2, the claim is proved. ■

Theorem 12: Let \bar{T}_{opt} be the optimal system time for the m -agent Limited Sensing DTRP over the class of spatially unbiased policies, and let $\mathcal{V}(\mathcal{Q})$ be a convex and equitable partition of \mathcal{Q} with respect to measure ψ where $\psi(q) = \sqrt{\varphi(q)}$ for all $q \in \mathcal{Q}$. If each agent $\ell \in I_m$ operates on its own region of dominance \mathcal{V}_ℓ under the UTTSP policy, then the system time satisfies

$$\frac{\bar{T}_{\text{UTTSP/ERD}}}{\bar{T}_{\text{opt}}} = 1 \quad \text{as } \lambda \rightarrow \infty. \quad (15)$$

Proof: We apply Theorem 8 to find the system time of a target, conditioned upon its location,

$$\mathbb{E}[T_{\text{UTTSP/ERD}}|q \in \mathcal{V}_\ell] = \frac{\beta^2}{2v^2} \lambda_\ell \left(\int_{\mathcal{V}_\ell} \sqrt{\varphi_\ell(q)} dq \right)^2.$$

Substituting $\varphi_\ell(q) = \varphi(q)/\varphi(\mathcal{V}_\ell)$ and $\lambda_\ell = \lambda\varphi(\mathcal{V}_\ell)$ we get

$$\mathbb{E}[T_{\text{UTTSP/ERD}}|q \in \mathcal{V}_\ell] = \frac{\beta^2 \lambda}{2v^2} \left(\int_{\mathcal{V}_\ell} \sqrt{\varphi(q)} dq \right)^2.$$

But the equitable partition with respect to $\psi(q) = \sqrt{\varphi(q)}$ implies that

$$\int_{\mathcal{V}_\ell} \sqrt{\varphi(q)} dq = \frac{1}{m} \int_{\mathcal{Q}} \sqrt{\varphi(q)} dq, \quad \forall \ell \in I_m,$$

and so

$$\mathbb{E}[T_{\text{UTTSP/ERD}}|q \in \mathcal{V}_\ell] = \frac{\beta^2 \lambda}{2m^2 v^2} \left(\int_{\mathcal{Q}} \sqrt{\varphi(q)} dq \right)^2.$$

Since the above is independent of the targets location, it is in fact the unconditioned system time. Moreover this implies that the strategy is spatially unbiased. Combining with Eq. (3), the claim is proved. ■

Theorem 13: Let φ be a piecewise uniform density, let \bar{T}_{opt} be the optimal system time for the m -agent Limited Sensing DTRP over the class of spatially biased policies, and let $\mathcal{V}(Q)$ be a convex and equitable partition of \mathcal{Q} with respect to measure ψ where $\psi(q) = \varphi(q)^{2/3}$ for all $q \in \mathcal{Q}$. If each agent $\ell \in I_m$ operates on its own region of dominance \mathcal{V}_ℓ under the BTTSP policy, then the system time satisfies

$$\frac{\bar{T}_{\text{BTTSP/ERD}}}{\bar{T}_{\text{opt}}} = 1 \quad \text{as } \lambda \rightarrow \infty. \quad (16)$$

Proof: We apply Theorem 9 to find the system time of a target, conditioned upon its location,

$$\mathbb{E}[T_{\text{BTTSP/ERD}}|q \in \mathcal{V}_\ell] = \frac{\beta^2}{2v^2} \lambda_\ell \left(\int_{\mathcal{V}_\ell} \varphi_\ell(q)^{2/3} dq \right)^3.$$

Substituting $\varphi_\ell(q) = \varphi(q)/\varphi(\mathcal{V}_\ell)$ and $\lambda_\ell = \lambda\varphi(\mathcal{V}_\ell)$ we get

$$\mathbb{E}[T_{\text{BTTSP/ERD}}|q \in \mathcal{V}_\ell] = \frac{\beta^2 \lambda}{2v^2} \frac{1}{\varphi(\mathcal{V}_\ell)} \left(\int_{\mathcal{V}_\ell} \varphi(q)^{2/3} dq \right)^3.$$

Unconditioning on the location of the target,

$$\begin{aligned} \bar{T}_{\text{BTTSP/ERD}} &= \sum_{\ell=1}^m \Pr[q \in \mathcal{V}_\ell] \cdot \mathbb{E}[T_{\text{BTTSP/ERD}}|q \in \mathcal{V}_\ell] \\ &= \sum_{\ell=1}^m \varphi(\mathcal{V}_\ell) \cdot \frac{\beta^2 \lambda}{2v^2} \frac{1}{\varphi(\mathcal{V}_\ell)} \left(\int_{\mathcal{V}_\ell} \varphi(q)^{2/3} dq \right)^3 \\ &= \frac{\beta^2 \lambda}{2v^2} \sum_{\ell=1}^m \left(\int_{\mathcal{V}_\ell} \varphi(q)^{2/3} dq \right)^3. \end{aligned}$$

But the equitable partition with respect to $\psi(q) = \varphi(q)^{2/3}$ implies that

$$\int_{\mathcal{V}_\ell} \varphi(q)^{2/3} dq = \frac{1}{m} \int_{\mathcal{Q}} \varphi(q)^{2/3} dq, \quad \forall \ell \in I_m,$$

and so

$$\bar{T}_{\text{BTTSP/ERD}} = \frac{\beta^2 \lambda}{2m^2 v^2} \left(\int_{\mathcal{Q}} \varphi(q)^{2/3} dq \right)^3.$$

Combining with Eq. (4), the claim is proved. ■

In summary, we have offered four policies, each of which performs optimally for the four cases studied: small sensing-range (spatially biased and unbiased) and heavy load (spatially biased and unbiased). In addition, we have offered a method by which to adapt the four policies to a multi-vehicle setting, retaining optimality, with minimal communication or collaboration. In particular, the agents partition the environment into regions of dominance, and each vehicle executes the single-vehicle policy on its own region, ignoring all others. However, the nature of the partition varies in the different cases addressed. Each scenario requires regions of dominance

TABLE I
OPTIMAL SYSTEM TIME FOR THE DTRP WITH LIMITED SENSING.

	spatially unbiased	spatially biased
$r \rightarrow 0^+$	$\frac{\Lambda}{4mvr}$	$\frac{1}{4mvr} \left(\int \sqrt{\varphi} dq \right)^2$
$\lambda \rightarrow \infty$	$\frac{\beta^2 \lambda^2}{2m^2 v^2} \left(\int \sqrt{\varphi} dq \right)^2$	$\frac{\beta^2 \lambda^2}{2m^2 v^2} \left(\int \varphi^{2/3} dq \right)^3$

equitable with respect to a measure appearing in the optimal system time of the corresponding single-vehicle case. In the spatially unbiased small sensing-range case, the regions of dominance are equitable with respect to area. Interestingly, the spatially biased small sensing-range case, and the spatially unbiased heavy load case both require regions of dominance equitable with respect to measure ψ where $\psi(q) = \sqrt{\varphi(q)}$ for all $q \in \mathcal{Q}$. Finally, the spatially biased heavy load case requires regions of dominance equitable with respect to measure ψ where $\psi(q) = \varphi(q)^{2/3}$ for all $q \in \mathcal{Q}$.

We have shown that in heavy load, the limited information gathering capabilities of the agents have no effect on their achievable performance. We suggest the intuitive explanation that in heavy load, the environment is dense with targets, and so the searching component added to the DTRP is non-existent. On the other hand, for small sensing range, the lack of information gathering capability detracts from the achievable performance of the system. However, the presented policies for these cases were still proved optimal through new lower bounds related to the searching capability of the agents and the necessary structure of any stabilizing policy.

VI. CONCLUSIONS

We have addressed a multi-agent problem with information constraints we call the DTRP with limited sensing. Our analysis yields precise characterizations of the system time, and the parameters describing the capabilities and limitations of the agents and the nature of the environment appear (or don't appear) in these expressions, giving insight into how the parameters affect (or don't affect) the efficiency of the system. We summarize the characterizations of the optimal system times for the four cases studied in Table VI.

In terms of methods and approaches, we note the path taken in the small-sensor case. In order to place lower bounds on the achievable performance of any algorithm, we carefully relax constraints to arrive at a convex optimization problem whose solution offers insight into the structure of the optimal algorithm. Using this structure as guidance, we design a provably optimal algorithm. Moreover, we have made use of results from the mature fields of combinatorial optimization and probability theory.

Another interesting result is that the limited sensing capability has no impact on performance when the target arrival-rate is high. In other words, this lack of global information does not hinder an agent from efficient routing choices. This result assumes knowledge of prior statistics on the global environment. Perhaps through proper mechanism design, a game-theoretic approach [18], [21] might integrate the learning of the global structure of the environment and the adaptation of policy choices. Consider a team of agents with limited

sensing and no communication [45] operating in a common environment. From the perspective of an individual agent, there might not be any observed difference between i) a region with low target-arrival rate, and ii) a region with high target-arrival rate that is frequently serviced by other agents. Does this difference matter? In either case, the agent should search for target-rich regions in order to maximize both its own utility (target-servicing rate) and the value it is adding to the multi-agent system.

REFERENCES

- [1] H. Psaraftis, "Dynamic vehicle routing problems," in *Vehicle Routing: Methods and Studies* (B. Golden and A. Assad, eds.), Studies in Management Science and Systems, Elsevier, 1988.
- [2] D. J. Bertsimas and G. J. van Ryzin, "A stochastic and dynamic vehicle routing problem in the Euclidean plane," *Operations Research*, vol. 39, pp. 601–615, Jul-Aug 1991.
- [3] D. J. Bertsimas and G. J. van Ryzin, "Stochastic and dynamic vehicle routing with general demand and interarrival time distributions," *Advances in Applied Probability*, vol. 25, no. 4, pp. 947–978, 1993.
- [4] K. Savla, E. Frazzoli, and F. Bullo, "Traveling salesperson problems for the dubins vehicle," *IEEE Trans. on Automatic Control*, vol. 53, no. 9, 2008.
- [5] J. J. Enright, K. Savla, F. Bullo, and E. Frazzoli, "Stochastic and dynamic routing problems for multiple uninhabited aerial vehicles," *AIAA Journal of Guidance, Control, and Dynamics*, vol. 32, no. 4, 2009.
- [6] J. J. Enright, K. Savla, and E. Frazzoli, "Coverage control for teams of nonholonomic agents," in *IEEE Conf. on Decision and Control*, (Cancun, Mexico), December 2008.
- [7] H. Choset, "Coverage of known spaces: the boustrophedon cellular decomposition," *Autonomous Robots*, vol. 9, no. 3, pp. 247–253, 2000.
- [8] Y. Guo and M. Balakrishnan, "Complete coverage control for nonholonomic mobile robots in dynamic environments," in *Int'l Coinference on Robotics and Automation*, (Orlando, FL), pp. 1704–1709, 2006.
- [9] C. Schumacher, P. R. Chandler, S. J. Rasmussen, and D. Waker, "Task allocation for wide area search munitions with variable path length," in *American Control Conference*, pp. 3472–3477, Denver, CO, 2003.
- [10] Y. Liu, J. Cruz, and A. G. Sparks, "Coordinated networked uninhabited aer vehicles for persistent area denial," in *IEEE Conf. on Decision and Control*, (Paradise Island, Bahamas), pp. 3351–3356, 2004.
- [11] J. Cortés, S. Martínez, T. Karatas, and F. Bullo, "Coverage control for mobile sensing networks," *IEEE Transactions on Robotics and Automation*, vol. 20, no. 2, pp. 243–255, 2004.
- [12] B. Moore and K. Passino, "Distributed balancing of aavs for uniform surveillance coverage," in *IEEE Conference on Decision and Control*, 2005.
- [13] R. W. Beard, T. W. McLain, M. A. Goodrich, and E. P. Anderson, "Coordinated target assignment and intercept for unmanned air vehicles," *IEEE Trans. on Robotics and Automation*, vol. 18, no. 6, pp. 911–922, 2002.
- [14] M. Earl and R. D'Andrea, "Iterative milp methods for vehicle control problems," *IEEE Trans. on Robotics*, vol. 21, pp. 1158–1167, December 2005.
- [15] W. Li and C. Cassandras, "A cooperative receding horizon controller for multivehicle uncertain environments," *IEEE Trans. on Automatic Control*, vol. 51, no. 2, pp. 242–257, 2006.
- [16] X. Ma and D. A. Castanon, "Receding horizon planning for dubins traveling salesman problems," in *IEEE Conference on Decision and Control*, (San Diego, CA), Dec 2006.
- [17] R. Murphey, "Target-based weapon target assignment problems," in *Non-linear Assignment Problems: Algorithms and Applications* (P. Pardalos and L. Pitsoulis, eds.), pp. 39–53, Kluwer Academic Publisher, 1999.
- [18] G. Arslan, J. Marden, and J. Shamma, "Autonomous vehicle-target assignment: a game theoretic formulation," *ASME Journal of Dynamic Systems, Measurement and Control*, vol. 129, pp. 584–596, September 2007.
- [19] B. Bethke, M. Valenti, and J. P. How, "Uav task assignment," *IEEE Robotics and Automation Magazine*, March 2008.
- [20] M. Alighanbari, L. F. Bertuccelli, and J. P. How, "A robust approach to the UAV task assignment problem," in *IEEE Conf. on Decision and Control*, (San Diego, CA), December 2006.
- [21] J. Shamma, *Cooperative Control of Distributed Multi-Agent Systems*. John Wiley & Sons, 2007.
- [22] M. M. Zavlanos and G. J. Pappas, "Dynamic assignment in distributed motion planning with limited information," in *American Control Conference*, (New York, NY), July 2007.
- [23] J. Ousingsawat and M. E. Campbell, "Optimal cooperative reconnaissance using multiple vehicles," *AIAA Journal of Guidance, Control, and Dynamics*, vol. 30, pp. 122–132, January-February 2007.
- [24] L. F. Bertuccelli, M. Alighanbari, and J. P. How, "Robust planning for coupled cooperative UAV missions," in *IEEE Conf. on Decision and Control*, (Atlantis, Paradise Island, Bahamas), December 2004.
- [25] B. A. Kish, D. R. Jacques, and M. Pachter, "Optimal control of sensor threshold for autonomous wide-area-search munitions," *AIAA Journal of Guidance, Control, and Dynamics*, vol. 30, pp. 1239–1248, September-October 2007.
- [26] D. Song, C. Y. Kim, and J. Yi, "Stochastic modeling of the expected time to search for an intermittent signal source under a limited sensing range," in *Proceedings of Robotics: Science and Systems*, (Zaragoza, Spain), June 2010.
- [27] G. Mathew and I. Mezic, "Spectral multiscale coverage: A uniform coverage algorithm for mobile sensor networks," in *Proceedings of the 48th IEEE Control and Decision Conference*, (Shanghai, China), pp. 7872–7877, December 2009.
- [28] G. Cannata and A. Sgorbissa, "A minimalist algorithm for multirobot continuous coverage," *IEEE Trans. on Robotics*, vol. 27, no. 2, 2011.
- [29] J. J. Enright and E. Frazzoli, "Cooperative UAV routing with limited sensor range," in *AIAA Conf. on Guidance, Navigation, and Control*, (Keystone, CO), August 2006.
- [30] V. A. Huynh, J. J. Enright, and E. Frazzoli, "Persistent patrol with limited-range on-board sensors," in *IEEE Conf. on Decision and Control*, 2010.
- [31] J. Beardwood, J. Halton, and J. Hammersley, "The shortest path through many points," *Proceedings of the Cambridge Philosophy Society (Mathematical and Physical Sciences)*, vol. 55, no. 4, pp. 299–327, 1959.
- [32] J. M. Steele, "Probabilistic and worst case analyses of classical problems of combinatorial optimization in Euclidean space," *Mathematics of Operations Research*, vol. 15, no. 4, p. 749, 1990.
- [33] G. Percus and O. C. Martin, "Finite size and dimensional dependence of the Euclidean traveling salesman problem," *Physical Review Letters*, vol. 76, no. 8, pp. 1188–1191, 1996.
- [34] D. S. Johnson, L. A. McGeoch, and E. E. Rothberg, "Asymptotic experimental analysis for the held-karp traveling salesman bound," in *Proc. 7th Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 341–350, 1996.
- [35] R. Larson and A. Odoni, *Urban Operations Research*. Englewood Cliffs, NJ: Prentice Hall, 1981.
- [36] S. Lin and B. W. Kernighan, "An effective heuristic algorithm for the traveling-salesman problem," *Operations Research*, vol. 21, pp. 498–516, 1973.
- [37] N. Christofides, "Bounds for the travelling-salesman problem," *Operations Research*, vol. 20, pp. 1044–1056, 1972.
- [38] D. J. Bertsimas and G. J. van Ryzin, "Stochastic and dynamic vehicle routing in the Euclidean plane with multiple capacitated vehicles," *Operations Research*, vol. 41, pp. 60–76, Jan-Feb 1993.
- [39] J. Little, "A proof of the queueing formula $l = \lambda w$," *Operations Research*, vol. 9, pp. 383–387, 1961.
- [40] H. Xu, *Optimal Policies for Stochastic and Dynamic Vehicle Routing Problems*. PhD thesis, Dept. of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA, September 1994.
- [41] R. G. Gallager, *Discrete Stochastic Processes*. Kluwer Academic Publishers, 1996.
- [42] D. P. Bertsekas, *Nonlinear Programming*. Belmont, MA: Athena Scientific, 1995.
- [43] R. Kershner, "The number of circles covering a set," *American Journal of Mathematics*, vol. 61, pp. 665–671, July 1939.
- [44] M. Pavone, E. Frazzoli, and F. Bullo, "Distributed policies for equitable partitioning: Theory and applications," in *IEEE Conference on Decision and Control*, (Cancun, Mexico), Dec 2008.
- [45] A. Arsie and E. Frazzoli, "Efficient routing algorithms for multiple vehicles with no communications," *IEEE Trans. on Automatic Control*, vol. 54, no. 10, 2009.